AMERICAN MUSEUM
Novitates

PUBLISHED BY THE AMERICAN MUSEUM OF NATURAL HISTORY
CENTRAL PARK WEST AT 79TH STREET, NEW YORK, NY 10024
Number 3652, 10 pp., 2 figures June 25, 2009

Mitochondrial Intergenic Spacer in Fairy Basslets (Serranidae: Anthiinae) and the Simultaneous Analysis of Nucleotide and Rearrangement Data

WM. LEO SMITH,[1,2] KATHLEEN R. SMITH,[2] AND WARD C. WHEELER[3]

ABSTRACT

We present the results of a study that implements a recently developed phylogenetic algorithm that combines fixed-states nucleotide optimization with breakpoint analysis to identify and examine the evolution of a mitochondrial intergenic spacer between the $\mathrm{tRNA}^{\mathrm{Val}}$ and 16S rRNA loci in a clade of fairy basslets (Serranidae: Anthiinae). The results of the analysis indicate that this spacer evolved once and that it may be increasing in size through evolutionary time. The resulting molecular hypothesis corroborates much of the previous morphological phylogenetic work.

INTRODUCTION

During recent investigations focusing on the relationships of "mail-cheeked" fishes and sea basses (Smith and Wheeler, 2004, 2006; Smith and Craig, 2007), we discovered a mitochondrial intergenic spacer in a clade of anthiine sea basses. After surveying the analytical literature, we were concerned about the subjectivity of the methods that have been previously used to identify and analyze novel genetic elements and genomic rearrangement data. Therefore, we have used a recently developed genomic algorithm (Wheeler, 2007) to examine the evolution of this anthiine intergenic spacer and the relationships within anthiine serranids. We then discuss how this method can be expanded beyond this comparatively simple example to simultaneously analyze genomic and sequence data in a phylogenetic framework. During the last 20 years, evolutionary biologists have made

---

[1] Department of Zoology, Field Museum of Natural History, 1400 South Lake Shore Drive, Chicago, IL 60605, USA (lsmith@fieldmuseum.org).
[2] Department of Ichthyology, American Museum of Natural History, Central Park West at 79th Street New York, NY 10024, USA (ktsmith@amnh.org).
[3] Division of Invertebrate Zoology, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024, USA (wheeler@amnh.org).

Copyright © American Museum of Natural History 2009 ISSN 0003-0082

tremendous advances by combining innovations in automated DNA sequencing and computational biology. This progress has provided the data necessary to test evolutionary hypotheses using complete genomic sequences. Animal mitochondrial genomes (mitogenomes) represent the vast majority of complete genomic data. As the number of sequenced mitogenomes increases, so does our understanding of their organization and function. These studies confirm previous suggestions that animal mitogenomes are largely conserved in terms of gene order and content (Boore and Brown, 1998; Miya et al., 2005), but they also challenge previous assumptions of mitochondrial genome evolution (e.g., the lack of recombination; Piganeau et al., 2004). Although most sequenced animal mitogenomes (primarily vertebrate) share a particular gene order, rearrangements and other transformations have been described in a variety of animal groups (e.g., amphibians, squamates, teleosts, hymenopterans [Macey et al., 1997; Dowton et al., 2003; Miya et al., 2005]). These transformations include dramatic reductions in genomic size in chaetognaths (Helfenbein et al., 2004) to increases in size ranging from small tandem repeats (Moritz et al., 1987) to large-scale duplications (McKnight and Shaffer, 1997). The consequences of these duplications range from multiple functional copies of loci undergoing concerted evolution (Kumazawa et al., 1996) to rearrangements (Macey et al., 1997; Dowton et al., 2003; Miya et al., 2005) to intergenic spacers (McKnight and Shaffer, 1997; Bakke et al., 1999; Mabuchi et al., 2004). Many authors have argued that the assumed uniqueness and rarity of these genome-level transformations make them more reliable than nucleotide transformations for inferring deeper phylogenetic relationships (Moritz et al., 1987; Curole and Kocher, 1999). However, as additional mitogenomes are published, it is clear that these transformations are more common than once believed (Curole and Kocher, 1999; San Mauro et al., 2006), and, more importantly, it is striking how frequently these rearrangements occur in "hotspots" between transfer RNAs, often evolving repeatedly within large clades (Dowton et al., 2003; San Mauro et al., 2006). It seems clear that as the number of sequenced mitogenomes passes 2000, there will be continued discovery of novel and convergent rearrangements; these mitogenomic rearrangements foreshadow the complications that will be encountered as researchers begin to sequence the larger prokaryotic genomes and eukaryotic plastid and nuclear genomes. The realization that these rare mitogenomic rearrangements may arise through convergent evolution (Mindell et al., 1998; Curole and Kocher, 1999; Dowton et al., 2003) or be complicated by recombination (Piganeau et al., 2004) challenges the assumptions upon which previous analyses of these genomic characters were based. Most previous studies made a priori assumptions of locus homology based on sequence similarity, perceived secondary structure, or tRNA anticodon sequences (Dowton et al., 2003; Mabuchi et al., 2004; Miya et al., 2005). None of these methods use a historical, evolutionary framework for inferring or testing homology. Investigators using these nonphylogenetic approaches for identifying homologous loci will be misled by convergent evolution. For example, Rawlings et al. (2003) showed that $\mathrm{tRNA^{Leu}}$ had convergently altered its anticodon sequence at least seven times within metazoans. Therefore, it is clear that homology assessment should be explicitly tested—not assumed—in a comparative, evolutionary framework using all available evidence. When these scenarios are not quantitatively tested, they are prone to investigator bias or suboptimal explanations. This homology-assessment problem will only be exacerbated as we move from the simple animal mitogenomes to the larger prokaryotic genomes and eukaryotic plastid and nuclear genomes. Sankoff and Blanchette (1998), recognizing the need for explicit methods to infer phylogenies using genomic rearrangement data, developed breakpoint analysis, which infers phylogenetic relationships from locus order data. Unfortunately, breakpoint analysis ignores sequence variation and requires that locus homology be established prior to the analysis. As mentioned above, this locus homology assessment is difficult for all but the simplest cases. Recently, Wheeler (2007) described and implemented an algorithm in POY (Wheeler et al., 2003) that combines a breakpoint algorithm with a fixed-states nucleotide algorithm (Wheeler, 1999) to optimize chromosomal and nucleotide data simultaneously for inferring locus and nucleotide homology during phylogenetic analysis. This method does not assume locus homology or nucleotide alignment prior to analysis; it requires only that the boundaries of the loci be assigned (e.g., using start and stop codons for protein-coding sequences). Using evidence from the sequence data, this algorithm will infer homologous loci and simultaneously analyze both the nucleotide and genomic variation to recover the optimal phylogenetic relationships and hypothesize both nucleotide and locus homology in an evolutionary framework. Furthermore, this method is an improvement over current similarity-based or function-based approaches because it is automated, internally consistent, repeatable, explicit, does not presuppose a mechanism for resulting rearrangements, allows evaluation of competing hypotheses, and is grounded in a phylogenetic framework.

Herein, we use this new method to examine the evolution of this novel seabass intergenic spacer. Typically, vertebrate mitogenomes have two large intergenic spacers: the origin of light-strand replication, which has been lost in some snakes, crocodiles, and birds (Macey et al., 1997), and the control region (Fernández-Silva et al., 2003). Additional intergenic spacers are rare in vertebrate mitogenomes, but have been described between tRNA$^{\mathrm{Thr}}$ and tRNA$^{\mathrm{Pro}}$ in a variety of vertebrate clades, e.g., *Xenopus* (Roe et al., 1985), *Struthio* (Härlid et al., 1997), ambystomatid salamanders (McKnight and Shaffer, 1997), and gadiform fishes (Bakke et al., 1999). McKnight and Shaffer (1997) showed that the salamander intergenic spacer had the highest substitution rate among sequenced regions of the salamander mitogenome (including the control region). Furthermore, these authors noted that the persistence of this intergenic spacer contradicted Rand's (1993, 2001) hypotheses that mitogenomic size is minimized over time. Additional mitogenomic modifications and rearrangements have also been described and discussed by Mueller and Boore (2005) for other salamander taxa. The intergenic spacer described in this study is restricted to a subset of sea basses in the subfamily Anthiinae. The relationships of the sea basses (Serranidae and Epinephelidae) have received substantial attention (Johnson, 1983; Baldwin and Smith, 1998; Craig and Hastings, 2007; Smith and Craig, 2007), but the only explicit phylogenetic study, to date, of the Anthiinae is that of Baldwin (1990), who examined the relationships among American anthiines. Here, we report and characterize a vertebrate intergenic spacer discovered in H-strand transcription unit one (H1). Although comparatively rare, an H1 rearrangement has been noted in some toadfishes (Miya et al., 2005). The H1 includes tRNA$^{\mathrm{Phe}}$, tRNA$^{\mathrm{Val}}$, and both ribosomal RNAs (Fernández-Silva et al., 2003). Within H1, this anthiine intergenic spacer evolved between tRNA$^{\mathrm{Val}}$ and the 16S large ribosomal subunit (fig. 1). The remaining 24 heavy-strand loci are transcribed in H-strand transcription unit two (H2), along with some transcriptional overlap with H1. Given the smaller size and 20-fold greater transcription rate of H1 (Fernández-Silva et al., 2003), this first described H1 intergenic spacer may affect its posttranscriptional processing.

MATERIALS AND METHODS

ACQUISITION OF NUCLEOTIDE SEQUENCES AND TAXON SAMPLING

Fish tissues were preserved in 80% ethanol or were used fresh prior to extraction of DNA. Genomic DNA was extracted from muscle using a DNeasy Tissue Extraction Kit (Qiagen) following the manufacturer's protocol. Polymerase chain reaction (PCR) was used to amplify three overlapping fragments crossing three rDNA fragments (12S, tRNA$^{\mathrm{Val}}$, and 16S) and the intergenic spacer, when present (fig. 1). Multiple overlapping primer pairs were used to ensure that we were amplifying the intergenic spacer region in the mitochondrial genome instead of amplifying a nuclear pseudogene. Double-stranded amplifications were performed in a 25 μL volume containing one Ready-To-Go PCR bead (Amersham Biosciences), 1.25 μL of each primer (10 μM) and 2–5 μL of DNA. To amplify and sequence the tRNA$^{\mathrm{Val}}$, intergenic spacer (when present), and 5′ end of 16S fragment, the primers 12SL13-L 5′-TTA-

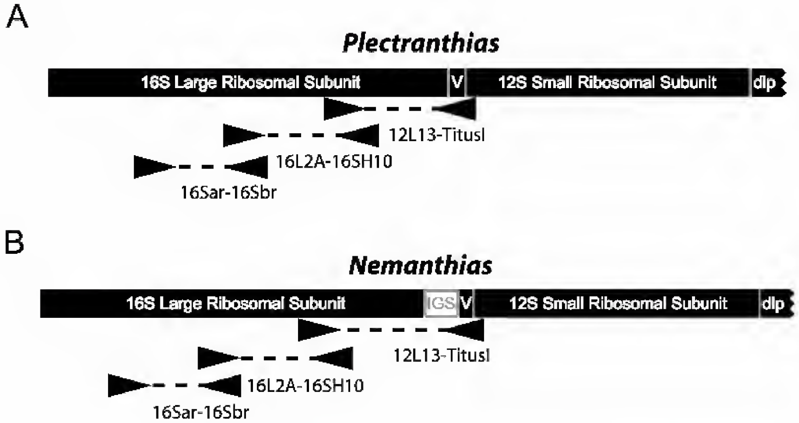

A
Plectranthias
16S Large Ribosomal Subunit
V
12S Small Ribosomal Subunit
dlp
12L13-TitusI
16L2A-16SH10
16Sar-16Sbr
B
Nemanthias
16S Large Ribosomal Subunit
IGS
V
12S Small Ribosomal Subunit
dlp
12L13-TitusI
16L2A-16SH10
16Sar-16Sbr

Fig. 1. Illustration of partial H1 in anthiine sea basses that (A) lack or (B) have evolved the intergenic spacer (IGS = intergenic spacer, V = tRNA^{Val}) with overlapping primer pairs used to amplify DNA mapped.

GAAGAGGCAAGTCGTAACATGGTA-3′ and TitusI-H 5′-GGTGGCTGCTTTTAGG-CC-3′ (Titus, 1992; Feller and Hedges, 1998) were used. To amplify the middle section of the 16S fragment, the primers 16SL2A-L 5′-C-CAAACGAGCCTAGTGATAGCTGGTT-3′ and 16SH10-H 5′-TGATTACGCTACCTTT-GCACGGT-3′ (Hedges, 1994) were used. To amplify the final section of the 16S fragment, the primers 16S ar-L 5′-CGCCTGTTTATC-AAAAACAT-3′ and 16S br-H 5′-CCGG-TCTGAACTCAGATCACGT-3′ (Kocher et al., 1989; Palumbi, 1996) were used. Amplifications for all fragments were carried out in 36 cycles using the following temperature profile: initial denaturation for 360 sec. at 94° C, denaturation for 60 sec. at 94° C, annealing for 60 sec. at 48° C, and extension for 75 sec. at 72° C, with an additional terminal extension at 72° C for 360 sec. The double-stranded amplification products were desalted and concentrated using AMPure (Agencourt® Bioscience Corporation) following the manufacturer's protocol. Both strands of the purified PCR fragments were used as templates and directly cycle-sequenced using the original amplification primers and Prism Dye Terminator Reaction Kit v1.1 (ABI). The sequencing reactions were cleaned and desalted using CleanSEQ (Agencourt® Bioscience Corporation) following the manufacturer's protocol. The nucleotides were sequenced on an ABI 3730×l automated DNA sequencer. Contigs were built in Sequencher (GeneCodes) using DNA sequences from the complementary heavy and light strands. Sequences were edited in Sequencher and Bioedit (Hall, 1999). All novel sequences were deposited in GenBank under accession numbers FJ548764–FJ548784. The remaining three sequences (*Myripristis berndti* [NC 003189], *Caranx melampygus* [NC 004406], and *Holanthias chrysostictus* [AY141436]) were taken from GenBank and were derived from studies by Miya et al. (2005) and Chen et al. (2003). Mitochondrial DNA sequences were analyzed for 18 serranid and epinephelid species and three outgroups (*Myripristis*, *Caranx*, and *Acanthistius*). We focused our taxon sampling on representatives of the smaller fairy basslets (e.g., *Luzonichthys*, *Nemanthias*, *Pseudanthias*, *Tosana*) in the subfamily Anthiinae because these genera were initially found to posses the intergenic spacer in H1. We also included representatives of four additional anthiine genera (*Hemanthias*, *Holanthias*, *Plectranthias*, and *Pronotogrammus*), three serranine genera (*Diplectrum*, *Paralabrax*, and *Zalanthias*), two epinephelid genera (*Epinephelus* and *Grammistes*), and the former anthiine, now percoid, genus *Acanthistius* to examine the phylogenetic distribution of the intergenic spacer. Furthermore, we included four specimens of the species *Pseudanthias tuka* to assess haplotype diversity for the intergenic spacer region. All four of these specimens were procured at one time from the aquarium trade, so their collection localies are unknown.

ANALYSIS

For the phylogenetic analysis, 17,285 aligned base pairs (based on the implied alignment; Wheeler, 2003) were analyzed. This large number of base pairs resulted from our inclusion of the complete mitogenomes for two taxa (*Myripristis* and *Caranx*). The remaining terminals sequenced in this study included the sequenced regions of the $\mathrm{tRNA}^{\mathrm{Val}}$, intergenic spacer (when present), and 16S, encompassing 2180 aligned base pairs (fig. 1). These sequences were initially analyzed simultaneously under the optimality criterion of parsimony with nucleotide transformations (e.g., transversions, transitions, indels) equally weighted with a cost of one, a breakpoint cost (the cost for separating adjacent loci) of 100, a locus gap cost (the constant fraction of the cost of locus origin-loss) of 200, and a locus-size gap cost (the variable fraction of the cost of locus origin-loss based on the size of the locus) of one. This analysis was repeated with a variety of other plausible weighting schemes with weights from one to 500 for the various breakpoint parameters. The limited changes in results that we recovered will be discussed below, but all inferences must be based on the explicit costs.

This analysis was conducted in POY (Wheeler et al., 2003, 2006) and run on the American Museum of Natural History parallel computing cluster. The analysis began with 500 random addition sequences that were improved with TBR branch swapping and 150 parsimony ratchet replicates (Nixon, 1999). The results of these analyses were submitted to a final round of tree fusing (Goloboff, 1999a) and TBR branch swapping using fixed-states character optimization (Wheeler, 1999) and breakpoint analysis (Sankoff and Blanchette, 1998). To estimate the robustness of the clades recovered in the phylogenetic hypotheses, Bremer supports (Bremer, 1994) were calculated using TreeRot (Sorenson, 1999) in conjunction with PAUP* 4.0b10 (Swofford, 2002), and jackknife resampling analyses were performed in NONA (Goloboff, 1999b; 1,000 replications, heuristic searches, 10 random additions per replication via the WinClada interface [Nixon, 2002]) using the implied alignment, which does not include the breakpoint costs. In addition to the POY analysis, we manipulated and compared sequence data using ClustalX (Thompson et al., 1997) using default values, and we examined possible secondary structures of the intergenic spacer transcript using the program MFOLD (Zuker et al., 1999) using default values.

RESULTS

The POY analysis results in 15 equally most parsimonious trees (strict consensus shown in fig. 2). Each of the resulting optimal hypotheses has a cost of 24,036 weighted steps, which includes the traditional nucleotide transformation costs as well as POY's rearrangements costs. Based on this analysis, we hypothesize a single origin of the intergenic spacer in the ancestor of the anthiine genera *Pseudanthias*, *Luzonichthys*, *Tosana*, and *Nemanthias*. Because the analysis includes two complete mitogenomes, it explicitly tests whether the intergenic spacer evolved de novo (e.g., duplication, insertion) or is a rearrangement from elsewhere in the mitogenome. The analysis indicates that the intergenic spacer is a novel element because it was not aligned (i.e., inferred to be homologous) with any mitogenomic element in the analyses that included the entire ~2100 base-pair region that was sequenced. Subsequent POY and ClustalX analyses that compared the intergenic spacer sequences individually to both included mitogenomes consistently aligned it with $\mathrm{tRNA}^{\mathrm{Val}}$ despite the inclusion of all mitochondrial transfer RNAs for *Myripristis* and *Caranx*. This suggests that the intergenic spacer is a duplication of this adjacent tRNA. Using the resulting phylogenetic framework, we describe the general features of and discuss the evolution of this novel mitochondrial element as well as compare our phylogenetic results to those of Baldwin (1990). The $\mathrm{tRNA}^{\mathrm{Val}}$ and 16S fragments that flank the intergenic spacer, when present, align well with the homologous regions in the taxa that lack the intergenic spacer and were identified as homologs in the POY analysis. The intergenic spacers of all 10 terminals have unique sequences with high sequence variation (uncorrected "p" distances): 0.025–0.618 (mean: 0.409 ± σ: 0.149) between species.

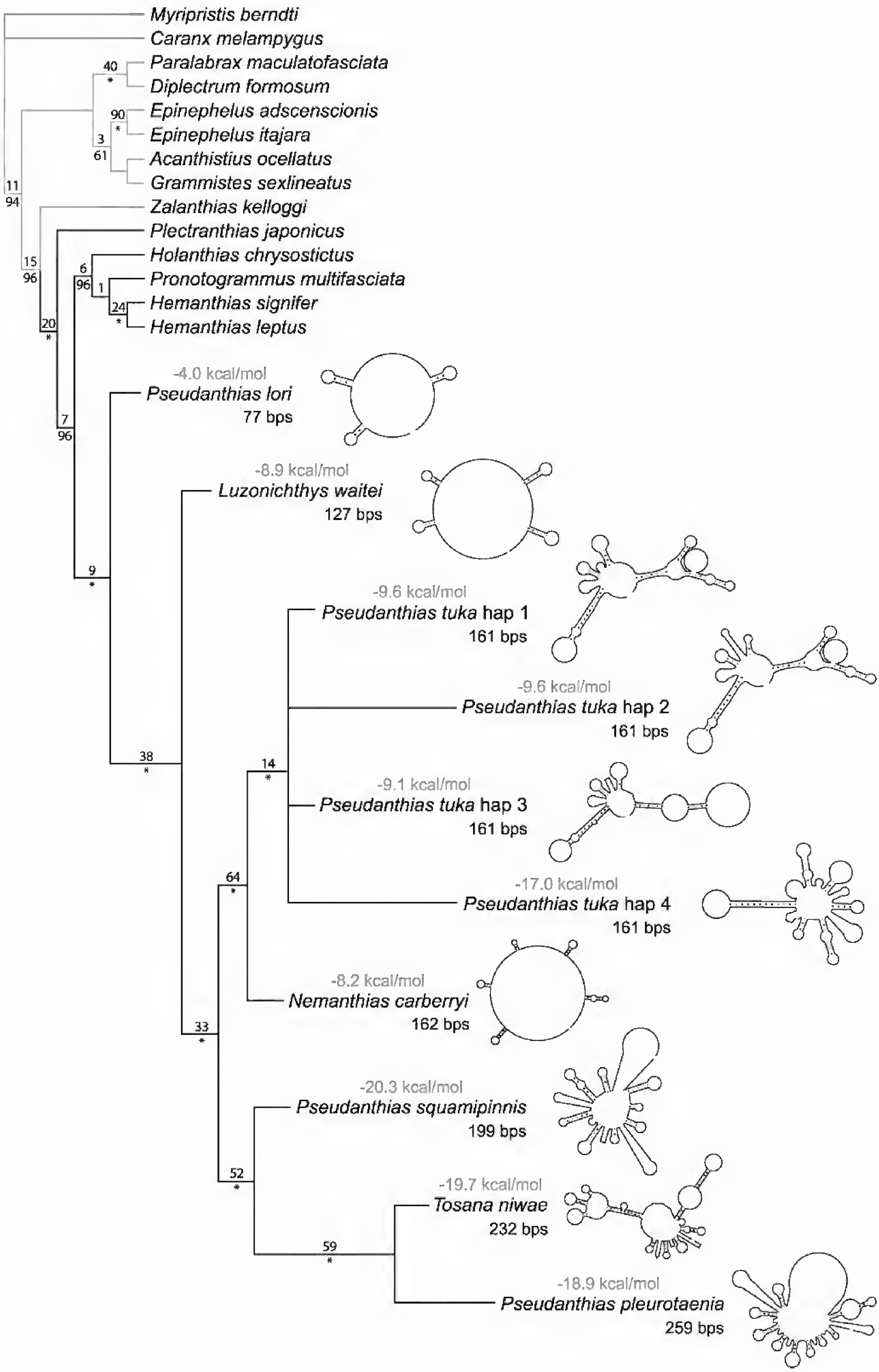
This variation is approximately three times that found in the flanking $\mathrm{tRNA}^{\mathrm{Val}}$ and 16S fragments for the same species: 0.010–0.181 (0.140 ± 0.041). All four *Pseudanthias tuka* haplotypes have distinct intergenic spacer sequences (two to nine variable sites) but identical $\mathrm{tRNA}^{\mathrm{Val}}$ and partial 16S sequences. Additionally, the intergenic spacer sequences are highly length variable, ranging from 77 to 259 bps (174 ± 62) between species (fig. 2). Despite the high variability, there is a region of ~25 base pairs at the 3′ end of the intergenic spacer sequences that is highly conserved among these 10 terminals.

DISCUSSION

As Wheeler (1995) described for nucleotide sequences, the results of this and all other numerical phylogenetic methods are affected by explicit assumptions (e.g., costs of transitions, transversions, indels, breakpoint cost, locus gap cost, locus-size gap cost) and implicit assumptions. Different costs for these various parameters may result in different hypotheses of relationships and evolutionary scenarios, as is seen with different weighting schemes in parsimony analyses or different evolutionary models in model-based phylogenetic approaches (i.e., MrBayes or maximum likelihood). Our resulting phylogeny was recovered in almost all breakpoint cost regimes that were examined. The specific breakpoint costs that we present for the calculations were chosen because they are roughly equivalent to the range of sizes for the intergenic spacer sequences. They were chosen so that the breakpoint costs would be equivalent to the number of independent substitutions and insertions required to explain the evolution of a single intergenic spacer sequence of observed length. In our diversity of analyses, the only topological changes that we found occurred when the breakpoint costs approached 1 (i.e., evidentially equivalent to a single substitution or indel). These results suggest that the intergenic spacer evolved independently (i.e., it was not homologous) in all species (except *Pseudanthias tuka* and *Nemanthias carberryi*) due to the extreme sequence variability in the intergenic spacer sequences found among species. These scenarios, as well as others, should be examined further as additional anthiine taxa are sequenced for the intergenic spacer region. Our discovery of this large, highly variable, yet persistent, intergenic spacer has implications for future studies examining the systematics and population genetics of sea basses, as well as for the evolutionary dynamics of the mitogenome. Our phylogenetic hypothesis recovered a monophyletic Anthiinae (sensu Smith and Craig, 2007) and corroborates Baldwin's (1990) and Smith and Craig's (2007) hypothesis that *Plectranthias* (sensu Eschmeyer, 1998) is not monophyletic and Baldwin's (1990) hypothesis that among included taxa, *Hemanthias*, *Holanthias*, and *Pronotogrammus* form a monophyletic group. Clearly, additional work on anthiine relationships is still desperately needed, but our study clearly traces and documents the evolution and persistence of this intergenic spacer in a diverse clade of anthiine serranids (presumably 65+ species in *Luzonichthys*, *Nemanthias*, *Pseudanthias*, and *Tosana* [Eschmeyer, 1998]). The persistence of this large intergenic spacer is as surprising as its increase in size through evolutionary time (fig. 2). Among the taxa sequenced in this study, the most basal taxon

→

Fig. 2. The strict consensus cladogram of the 15 most parsimonious trees. Anthiine serranids have black branches. The two-dimensional secondary structure of the transcript predicted by MFOLD for the intergenic spacer is adjacent to the taxon name, the free-energy calculation resulting from the secondary structure prediction of the intergenic spacer is listed above the taxon name, and the number of base pairs in the intergenic spacer are listed below the taxon name. Bremer supports are listed above the branches and jackknife values >50% are listed below the nodes (* = 100%). Nodes that lack Bremer support are recovered in the analysis that include breakpoints costs, but they are not recovered in the support value analyses based exclusively on the nucleotide transformations inferred from the implied alignment. Varied branch lengths do not represent amounts of change; they are varied solely to provide sufficient space for illustrating predicted secondary structures.


Myripristis berndti
Caranx melampygus
40
*
Paralabrax maculatofasciata
Diplectrum formosum
90
*
3
61
Epinephelus adscenscionis
Epinephelus itajara
Acanthistius ocellatus
Grammistes sexlineatus
11
94
Zalanthias kelloggi
15
96
Plectranthias japonicus
6
96
1
Holanthias chrysostictus
Pronotogrammus multifasciata
24
*
Hemanthias signifer
Hemanthias leptus
20
*
-4.0 kcal/mol
Pseudanthias lori
77 bps
7
96
-8.9 kcal/mol
Luzonichthys waitei
127 bps
9
*
-9.6 kcal/mol
Pseudanthias tuka hap 1
161 bps
-9.6 kcal/mol
Pseudanthias tuka hap 2
161 bps
38
*
14
*
-9.1 kcal/mol
Pseudanthias tuka hap 3
161 bps
64
*
-17.0 kcal/mol
Pseudanthias tuka hap 4
161 bps
-8.2 kcal/mol
Nemanthias carberryi
162 bps
33
*
-20.3 kcal/mol
Pseudanthias squamipinnis
199 bps
52
*
-19.7 kcal/mol
Tosana niwae
232 bps
59
*
-18.9 kcal/mol
Pseudanthias pleurotaenia
259 bps

with an intergenic spacer, *Pseudanthias lori*, has the smallest spacer (77 bp), roughly the size of a tRNA. At each successive node, there is an increase in the size of the intergenic spacer, culminating in more highly derived taxa with spacers as large as 259 bp (fig. 2). This trend of insertions within the intergenic spacer region is surprising considering current opinions about mitochondrial DNA evolution that argue for mitogenome size reduction through evolutionary time (Rand, 1993, 2001). Rand (1993, 2001) has argued that the lack of introns and intergenic spacers in animal mitochondrial DNA is due to strong purifying selection for small mitogenomic size (and thus rapid replication time). Given this argument, one would expect large, evolutionarily stable intergenic spacers to serve a function. None of the sea bass intergenic spacer sequences possess an open reading frame of realistic size, suggesting that they are not protein coding. As a further exploration of its function, we examined the predicted secondary structures of their RNA transcripts using a free energy minimization model in MFOLD. Although the intergenic spacer sequences often form complex two-dimensional stem and loop structures (fig. 2), these structures differ between species and haplotypes. Furthermore, they are not particularly stable, with free energies ranging from -4.0 to -20.3 kcal/mol, making it unlikely that they form functional secondary structures in organello. Additionally, we compared this tRNA^{Val}-16S region with more than 600 acanthomorph species from more than 250 families, and none of these taxa had a similar sequence between the tRNA^{Val} and 16S. Finally, we are unable to find any sequences similar to the intergenic spacer sequences by performing BLAST searches in GenBank. Therefore, it seems likely that the only functional role the intergenic spacer may play is in posttranscriptional processing. We speculate that this conserved 25-bp region is important for accurate cleavage at the 5′ end of the adjacent 16S rRNA, a function that would have to have been retained from the ancestral tRNA^{Val} following the hypothesized duplication to ensure proper posttranscriptional processing of the 16S rRNA. Ultimately, we agree with Curole and Kocher's (1999) assessment that "the possibility of convergent evolution of gene order can no longer be discounted." Therefore, we have analyzed our data using POY's new algorithm (Wheeler, 2007) that combines Sankoff and Blanchette's (1998) chromosomal rearrangement (breakpoint analysis) and Wheeler's (1999) fixed-states nucleotide optimizations to examine simultaneously the identity and evolution of this novel genomic locus as well as examine relationships among a subset of sea basses. This method for simultaneously combining chromosomal order and nucleotide data to assess locus homology (i.e., genomic annotation) and nucleotide alignment is preferred over other methods because it is automated, internally consistent, repeatable, explicit, does not presuppose a mechanism for resulting rearrangements, allows evaluation of competing hypotheses, and is grounded in a phylogenetic framework. Our results suggest a single evolution of an intergenic spacer in the highly transcribed H-strand transcription unit one, presumably through duplication and subsequent evolution of tRNA^{Val}. Finally, contrary to current mitochondrial genomic selection theory, this intergenic spacer is not only persistent, but seems to be increasing in size as it evolves.

ACKNOWLEDGMENTS

We would like to thank J. Smith (Los Alamos National Laboratory) and J. Faivovich, T. Grant, K. Pickett, J. Sparks, M. Stiassny, and K. Tang (all at or formerly at the American Museum of Natural History [AMNH]) for discussing aspects of this project with us. We are grateful to H. Endo (Kochi University), the Gahan Family, J. Leis and M. McGrouther (Australian Museum), Reef and Fin (Stamford, CT), and H. Walker (Scripps Institution of Oceanography) for providing specimens used in this study. This project was supported by funding from the AMNH Lerner-Gray Program for Marine Research, the NASA–Ames Fundamental Space Biology Program, the Field Museum of Natural History, and the National Science Foundation (DEB-0405246 and DEB-0732642).

REFERENCES

Bakke, I., G.F. Shields, and S. Johnansen. 1999. Sequence characterization of a unique intergenic spacer in Gadiformes mitochondrial DNA. Marine Biotechnology 1: 411–415.

Baldwin, C.C. 1990. Morphology of the larvae of American Anthiinae (Teleostei: Serranidae), with comments on relationships within the subfamily. Copeia 1990: 913–954.

Baldwin, C.C., and W.L. Smith. 1998. *Belonoperca pylei*, a new species of seabass (Teleostei: Serranidae: Epinephelinae: Diploprionini) from the Cook Islands with comments on relationships among diploprionins. Ichthyological Research 45: 325–339.

Boore, J.L., and W.M. Brown. 1998. Big trees from little genomes: mitochondrial gene order as a phylogenetic tool. Current Opinions in Genetics and Development 8: 668–674.

Bremer, K. 1994. Branch support and tree stability. Cladistics 10: 295–304.

Chen, W.J., C. Bonillo, and G. Lecointre. 2003. Repeatability of clades as a criterion of reliability: a case study for molecular phylogeny of Acanthomorpha (Teleostei) with larger number of taxa. Molecular Phylogenetics and Evolution. 26: 262–288.

Craig, M.T., and P.A. Hastings. 2007. A molecular phylogeny of the groupers of the subfamily Epinephelinae (Serranidae) with a revised classification of the Epinephelini. Ichthyological Research 54: 1–17.

Curole, J.P., and T.D. Kocher. 1999. Mitogenomics: digging deeper with complete mitochondrial genomes. Trends in Ecology and Evolution 14: 394–398.

Dowton, M., L.R. Castro, S.L. Campbell, S.D. Bargon, and A.D. Austin. 2003. Frequent mitochondrial gene rearrangements at the hymenopteran nad3–nad5 junction. Journal of Molecular Evolution 56: 517–526.

Eschmeyer, W.N. 1998. Catalog of fishes. San Francisco: California Academy of Sciences, 3 vols.

Feller, A.E., and S.B. Hedges. 1998. Molecular evidence for the early history of living amphibians. Molecular Phylogenetics and Evolution 9: 509–516.

Fernández-Silva, P., J.A. Enriquez, and J. Montoya. 2003. Replication and transcription of mammalian mitochondrial DNA. Experimental Physiology 88: 41–56.

Goloboff, P.A. 1999a. Analyzing large data sets in reasonable times: solutions for composite optima. Cladistics 15: 415–428.

Goloboff, P.A. 1999b. NONA. Version 2. Tucumán, Argentina: Published by author.

Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acid Symposium Series 41: 95–98.

Härlid, A., A. Janke, and U. Arnanson. 1997. The mtDNA sequence of the ostrich and the divergence between paleognathous and neognathous birds. Molecular Biology and Evolution 14: 754–761.

Hedges, S.B. 1994. Molecular evidence for the origin of birds. Proceedings of the National Academy of Sciences of the United States of America 91: 2621–2624.

Helfenbein, K.G., H.M. Fourcade, R.G. Vanjani, and J.L. Boore. 2004. The mitochondrial genome of *Paraspadella gotoi* is highly reduced and reveals that chaetognaths are a sister group to protostomes. Proceedings of the National Academy of Sciences of the United States of America 101: 10639–10643.

Johnson, G.D. 1983. *Niphon spinosus*: a primitive epinepheline serranid, with comments on the monophyly and intrarelationships of the Serranidae. Copeia 1983: 777–787.

Kocher, T.D., W.K. Thomas, A. Meyer, S.V. Edwards, S. Pääbo, F.X. Villablanca, and A.C. Wilson. 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. Proceedings of the National Academy of Sciences of the United States of America 86: 6196–6200.

Kumazawa, Y., H. Ota, M. Nishida, and T. Ozawa. 1996. Gene rearrangements in snake mitochondrial genomes: highly concerted evolution of control-region-like sequences duplicated and inserted into a tRNA gene cluster. Molecular Biology and Evolution 13: 1242–1254.

Mabuchi, K., M. Miya, T.P. Satoh, M.W. Westneat, and M. Nishida. 2004. Gene rearrangements and evolution of tRNA pseudogenes in the mitochondrial genome of the parrotfish (Teleostei: Perciformes: Scaridae). Journal of Molecular Evolution 59: 287–297.

Macey, J.R., A. Larson, N.B. Ananjeva, Z. Fang, and T.J. Papenfuss. 1997. Two novel gene orders and the role of light-strand replication in rearrangement of vertebrate mitochondrial genome. Molecular Biology and Evolution 14: 91–104.

McKnight, M.L., and H.B. Shaffer. 1997. Large, rapidly evolving intergenic spacers in the mitochondrial DNA of the salamander family Ambystomatidae (Amphibia: Caudata). Molecular Biology and Evolution 14: 1167–1176.

Mindell, D.P., M.D. Sorenson, and D.E. Dimcheff. 1998. Multiple independent origins of mitochondrial gene order in birds. Proceedings of the National Academy of Sciences of the United States of America 95: 10693–10697.

Miya, M., T.P. Satoh, and M. Nishida. 2005. The phylogenetic position of toadfishes (order

Batrachoidiformes) in the higher ray-finned fish as inferred from partitioned Bayesian analysis of 102 whole mitochondrial genome sequences. Biological Journal of the Linnean Society 85: 289–306.

Moritz, C., T.E. Dowling, and W.M. Brown. 1987. Evolution of animal mitochondrial DNA: relevance for population biology and systematics. Annual Reviews of Ecology and Systematics 18: 269–292.

Mueller, R.L., and J.L. Boore. 2005. Molecular mechanisms of extensive mitochondrial gene rearrangement in plethodontid salamanders. Molecular Biology and Evolution 22: 2104–2112.

Nixon, K.C. 1999. The parsimony ratchet, a new method for rapid parsimony analysis. Cladistics 15: 407–414.

Nixon, K.C. 2002. WinClada. Version 1.00.08. Ithaca, NY: Published by author.

Palumbi, S.R. 1996. Nucleic acids II: the polymerase chain reaction. *In* D.M. Hillis, C. Mortiz and B.K. Mable (editors), Molecular systematics. 2nd ed.: 205–247. Sunderland, MA: Sinauer Associates.

Piganeau, G., M. Gardner, and A. Eyre-Walker. 2004. A broad survey of recombination in animal mitochondria. Molecular Biology and Evolution 21: 2319–2325.

Rand, D.M. 1993. Endotherms, ectotherms, and mitochondrial genome-size variation. Journal of Molecular Evolution 37: 281–295.21.

Rand, D.M. 2001. The units of selection on mitochondrial DNA. Annual Reviews of Ecology and Systematics 32: 415–448.

Rawlings, T.A., T.M. Collins, and R. Bieler. 2003. Changing identities: tRNA duplication and remodeling within animal mitochondrial genomes. Proceedings of the National Academy of Sciences of the United States of America 100: 15700–15705.

Roe, B.A., D.P. Ma, R.K. Wilson, and J.F.H. Wong. 1985. The complete nucleotide sequence of the *Xenopus laevis* mitochondrial genome. Journal of Biological Chemistry 260: 9759–9774.

San Mauro, D., D.J. Gower, R. Zardoya, and M. Wilkinson. 2006. A hotspot of gene order rearrangement by tandem duplication and random loss in the vertebrate mitochondrial genome. Molecular Biology and Evolution 23: 227–234.

Sankoff, D., and M. Blanchette. 1998. Multiple genome rearrangement and breakpoint phylogeny. Journal of Computational Biology 5: 555–570.

Smith, W.L., and M.T. Craig. 2007. Casting the percomorph net widely: the importance of broad taxonomic sampling in the search for the placement of serranid and percid fishes. Copeia 2007: 35–55.

Smith, W.L., and W.C. Wheeler. 2004. Polyphyly of the mail-cheeked fishes (Teleostei: Scorpaeniformes): evidence from mitochondrial and nuclear sequence data. Molecular Phylogenetics and Evolution 32: 627–646.

Smith, W.L., and W.C. Wheeler. 2006. Venom evolution widespread in fishes: a road map for the bioprospecting of piscine venoms. Journal of Heredity 97: 206–217.

Sorenson, M.D. 1999. TreeRot. Version 2. Boston: Published by author.

Swofford, D.L. 2002. PAUP*. Version 4. Sunderland, MA: Sinauer Associates.

Thompson, J.D., T.J. Gibson, J. Plewniak, F. Jeanmougin, and D.G. Higgins. 1997. The ClustalX windows interface: flexible strategies for multiple sequence 22 alignment aided by quality analysis tools. Nucleic Acids Research 24: 4876–4882.

Titus, T.A. 1992. A phylogenetic analysis of the Desmognathinae (Caudata: Plethodontidae): evolutionary patterns inferred from mitochondrial DNA sequences. Ph.D. dissertation, University of Kansas, Lawerence, 122 pp.

Wheeler, W.C. 1995. Sequence alignment, parameter sensitivity, and the phylogenetic analysis of molecular data. Systematic Biology 44: 321–331.

Wheeler, W.C. 1999. Fixed character states and the optimization of molecular sequence data. Cladistics 15: 379–385.

Wheeler, W.C. 2003. Implied alignment: a synapomorphy-based multiple-sequence alignment method and its use in cladogram search. Cladistics 19: 261–268.

Wheeler, W.C. 2007. Chromosomal character optimization. Molecular Phylogenetics and Evolution 44: 1130–1140.

Wheeler, W.C., L. Aagesen, C.P. Arango, J. Faivovich, T. Grant, C. D'Haese, D. Janies, W.L. Smith, A. Varon, and G. Giribet. 2006. Dynamic homology and phylogenetic systematics: a unified approach using POY. New York: American Museum of Natural History.

Wheeler, W.C., D. Gladstein, and J. De Laet. 2003. POY. Version 3. New York: Published by the authors.

Zuker, M., D.H. Mathews, and D.H. Turner. 1999. Turner algorithms and thermodynamics for RNA secondary structure prediction. *In* J. Barciszewski and B.F.C. Clark (editors), A practical guide in RNA biochemistry and biotechnology: 11–43. New York: Kluwer Academic Publishers.

Complete lists of all issues of the *Novitates* and the *Bulletin* are available at World Wide Web site http://library.amnh.org/pubs. Inquire about ordering printed copies via e-mail from scipubs@amnh.org or via standard mail from: American Museum of Natural History, Library—Scientific Publications, Central Park West at 79th St., New York, NY 10024. TEL: (212) 769-5545. FAX: (212) 769-5009.

♾ This paper meets the requirements of ANSI/NISO Z39.48-1992 (Permanence of Paper).